eng

MDPI

Systematic Review

Machine Learning-Based Approaches for Early Detection and Risk Stratification of Deep Vein Thrombosis: A Systematic Review

Andre Axel Cadena Zepeda [1,†], Enrique Efrén García-Guerrero [1,*,†], Oscar Adrian Aguirre-Castro [1,†], Gilberto Manuel Galindo-Aldana [2,†], Reyes Juárez-Ramírez [3,†], Marco Antonio Gómez-Guzmán [1,†], Christian Raymond [4,†] and Everardo Inzunza-Gonzalez [1,*,†]

[1] Facultad de Ingeniería, Arquitectura y Diseño, Universidad Autónoma de Baja California, Carretera Transpeninsular Ensenada-Tijuana No. 3917, Ensenada 22860, B.C., Mexico; andre.cadena@uabc.edu.mx (A.A.C.-Z.); oscar.aguirre@uabc.edu.mx (O.A.A.-C.); marco.guzman60@uabc.edu.mx (M.A.G.-G.)
[2] Laboratory of Neuroscience and Cognition, Facultad de Ciencias Administrativas, Sociales e Ingeniería, Universidad Autónoma de Baja California, Carr. Est. No. 3 s/n Col. Gutierrez, Mexicali 21700, B.C., Mexico; gilberto.galindo.aldana@uabc.edu.mx
[3] Facultad de Ciencias Químicas e Ingeniería, Universidad Autónoma de Baja California, Calzada Universidad 14418, Parque Industrial Internacional Tijuana, Tijuana 22427, B.C., Mexico; reyesjua@uabc.edu.mx
[4] INSA Rennes, IRISA, 20 Avenue des Buttes de Coësmes, 35700 Rennes, France; christian.raymond@irisa.fr
* Correspondence: eegarcia@uabc.edu.mx (E.E.G.-G.); einzunza@uabc.edu.mx (E.I.-G.); Tel.: +52-646-175-0744 (E.E.G.-G. & E.I.-G.)
† These authors contributed equally to this work.

Abstract

Deep vein thrombosis is a condition associated with substantial morbidity and a high risk of pulmonary embolism, underscoring the need for rapid and reliable diagnostic solutions. Although machine learning and deep learning techniques are increasingly being applied for clinical decision support, comprehensive analyses of their contributions to early detection, risk prediction, and monitoring remain limited. Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 guidelines, we conducted a systematic search in ScienceDirect, IEEE Xplore, Scopus, and Web of Science for studies published between January 2014 and March 2025. Eligible studies applied machine learning or deep learning approaches for the early prediction, monitoring, or risk assessment of deep vein thrombosis, or described reference datasets for algorithm development. Two authors independently extracted data and evaluated methodological quality using the Quality Assessment of Diagnostic Accuracy Studies-2 framework. The included studies were categorized into four domains: Early prediction, monitoring, risk assessment, and reference datasets. In total, 66 studies met the inclusion criteria. Recent advances include deep learning-assisted ultrasound interpretation and real-time implementation of machine learning algorithms. While most studies demonstrated a low overall risk of bias, recurring limitations were identified in terms of patient selection, reporting practices, and validation strategies. Dataset harmonization and external validation were infrequently performed, and documentation of data provenance and class imbalance handling was inconsistent. Machine learning and deep learning approaches demonstrate considerable potential to accelerate accurate diagnoses and facilitate individualized risk stratification; however, their translation into routine practice requires standardized datasets, rigorous external validation, and integration into existing clinical workflows. This review consolidates a decade of research, links methodological quality to clinical applicability, and provides a task-oriented roadmap for advancing machine learning-enabled diagnostics and monitoring in the context of deep vein thrombosis.

check for updates

Academic Editor: Hisham Daoud

Received: 10 August 2025
Revised: 1 September 2025
Accepted: 11 September 2025
Published: 14 September 2025

Citation: Cadena Zepeda, A.A.; García-Guerrero, E.E.; Aguirre-Castro, O.A.; Galindo-Aldana, G.M.; Juárez-Ramírez, R.; Raymond, C.; Gómez-Guzmán, M.A.; Inzunza-Gonzalez, E. Machine Learning-Based Approaches for Early Detection and Risk Stratification of Deep Vein Thrombosis: A Systematic Review. *Eng* **2025**, *6*, 243. https://doi.org/10.3390/eng6090243

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Keywords: deep vein thrombosis (DVT) ; pulmonary embolism; compression ultrasonography; early detection; risk stratification; real-time monitoring; machine learning; deep learning; artificial intelligence

1. Introduction

Deep vein thrombosis (DVT) is recognized as the leading preventable cause of in-hospital mortality. This vascular condition and its associated complications pose a substantial clinical challenge, affecting multiple hospital departments and often arising as a secondary condition unrelated to the patient's primary diagnosis. Statistically, it ranks as the third leading cause of cardiovascular-related deaths, following acute myocardial infarction and stroke [1–4]. The rising incidence of DVT, including in pediatric populations, emphasizes its growing relevance to public health and the need for early diagnostic strategies. A major epidemiological study conducted across one-third of pediatric hospitals in the United States reported a 70% increase in DVT incidence between 2001 and 2007, rising from 38 to 58 cases per 10,000 children. This rise has been attributed to various aspects of modern life, including sedentary behavior and the increased use of invasive medical devices [5,6]. Consequently, early detection of DVT is critical for improving patient outcomes and reducing associated clinical risks.

In recent years, machine learning (ML) has emerged as a transformative tool in medicine, demonstrating significant potential to improve diagnostic workflows. Unlike conventional approaches, ML models can exploit large, heterogeneous datasets and reveal patterns that are not easily detected by clinicians. Countries such as the United States, Italy, the United Kingdom, Germany, and Canada have taken a leading role in both foundational and applied research involving ML-based disease detection. In the context of DVT, ML algorithms such as Decision Trees (DT), K-Nearest Neighbors (KNN), Multilayer Perceptron Neural Networks (MLP-NNs), Support Vector Machines (SVMs), and Random Forests (RFs) have shown promising results [3–8]. Several studies have suggested that ML models can surpass conventional diagnostic approaches in terms of both accuracy and operational efficiency. For example, deep learning techniques have been applied to compression ultrasound interpretation, thereby supporting non-specialists in the diagnostic process. These technological advances demonstrate the potential of ML to democratize diagnostic access and reduce reliance on expert personnel; however, their practical translation remains inconsistent.

Despite this progress, previous studies and reviews have often focused on specific datasets or algorithms, lacking a unified perspective on validation practices, real-world applicability, and generalizability. Existing works rarely provide a critical synthesis that compares model performance across diverse clinical contexts, which is essential for clinical adoption. Significant challenges persist, including dataset heterogeneity, limited external validation, and uncertainty regarding real-world performance. This fragmentation hinders the development of robust, scalable, and clinically validated ML solutions for DVT. Addressing this gap requires a review that not only catalogs the ML approaches applied to DVT but also critically contrasts methodologies, highlights limitations, and identifies research priorities.

This review therefore contributes by systematically evaluating ML-based strategies for the early detection, risk prediction, and monitoring of DVT, presenting a structured comparison of algorithms, datasets, and evaluation metrics. Emphasis is placed on analyzing the strengths and weaknesses of current approaches, providing actionable insights for future studies and clinical translation. By doing so, this work not only synthesizes the

available evidence but also highlights limitations in existing approaches, uncovers actionable research gaps, and provides recommendations for advancing clinically meaningful ML applications.

The remainder of this review is organized as follows: Section 2 details the methodology employed. Section 3 presents the findings of the review. Section 4 contextualizes these findings. Section 5 highlights emerging trends and future research directions. Section 6 addresses current limitations and clinical implications. Finally, Section 7 summarizes the main conclusions of this systematic review.

2. Methodology

A systematic literature search was conducted across ScienceDirect, IEEE Xplore, Scopus, and Web of Science to ensure comprehensive coverage of relevant studies. The following Boolean query was applied: ("Deep Vein Thrombosis" OR "DVT") AND ("Artificial Intelligence" OR "Deep Learning" OR "Machine Learning") AND ("Mobile application"). This combination of keywords was designed to maximize the retrieval of studies focusing on ML-based approaches to DVT detection and monitoring. Search terms were intentionally kept simple and precise, reflecting the most common terminology in the field. Duplicate records were identified and removed before the screening process.

To ensure methodological transparency and reproducibility, this review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines. A predefined protocol guided the search strategy, screening process, eligibility assessment, and data extraction. Multiple reviewers independently verified all decisions regarding article inclusion to minimize bias and maintain consistency. These measures strengthen the rigor and reliability of the review's findings.

2.1. Screening and Eligibility Results

A comprehensive multi-stage screening process was implemented to select studies that were both relevant to the objectives of this review and methodologically robust. Titles and abstracts were first examined to identify articles explicitly addressing ML techniques for DVT detection, prediction, or related clinical tasks. Articles passing this stage underwent a detailed full-text review, applying the predefined eligibility criteria summarized in Table 1.

Table 1. Eligibility criteria for study selection.

Criterion	Inclusion Criteria	Exclusion Criteria
Focus of Study	Studies explicitly reporting the application of ML techniques for DVT detection, prediction, or monitoring using clinical data, medical imaging (e.g., ultrasound, MRI), or electronic health records.	Studies not addressing ML-based approaches or not focusing on DVT diagnosis, prediction, or monitoring.
Type of Article	Original research articles providing substantial methodological contributions.	Literature reviews, editorials, commentaries, or isolated case reports.
Language of Publication	Articles published in English to ensure accurate interpretation and assessment consistency.	Articles published in other languages, excluded to minimize translation bias and maintain methodological rigor.

The first criterion required that studies explicitly apply ML techniques for DVT detection, risk prediction, or monitoring using clinical data, medical imaging (e.g., ultrasound, MRI), or electronic health records. The second criterion excluded literature reviews, editorials, commentaries, and isolated case reports, focusing solely on original research that offers substantive methodological contributions. The third criterion excluded articles published

in languages other than English. This restriction was applied to ensure accurate interpretation of clinical and technical terminology, avoid translation-related bias, and maintain consistency during quality assessment. Language restrictions are widely recognized as a practical approach in systematic reviews, particularly when translation resources are unavailable or insufficient to guarantee methodological rigor.

After applying these criteria, eligible studies were organized into four thematic categories:

- DVT prediction;
- DVT monitoring;
- DVT risk assessment;
- Reference data for algorithm development.

The article selection process was validated through consensus among reviewers and adhered strictly to the PRISMA 2020 guidelines to ensure transparency and reproducibility. The full selection workflow is illustrated in the PRISMA flow diagram in Figure 1.

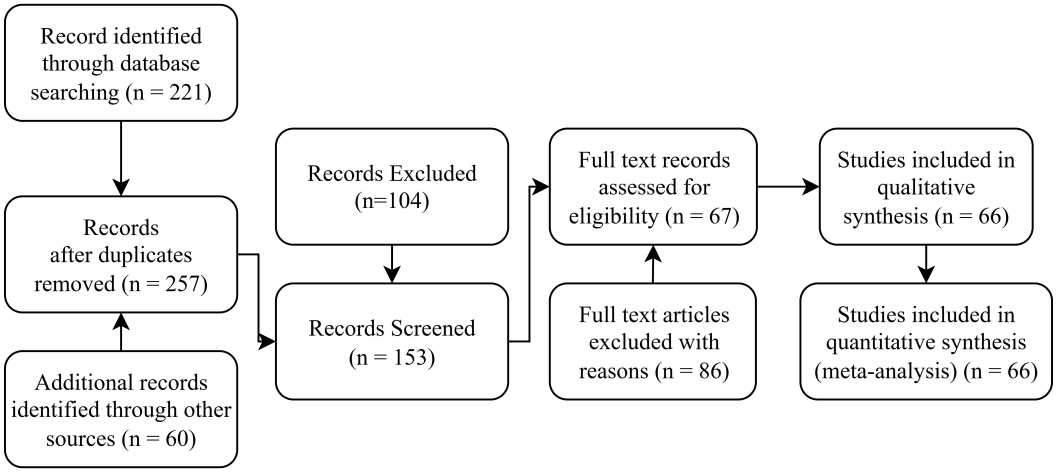

Figure 1. PRISMA 2020 flow diagram summarizing the multi-stage article selection process. The diagram shows the number of records identified, screened, and assessed for eligibility, and included in the final synthesis.

2.2. Reviewer Roles

The selection and evaluation of studies were conducted independently by two reviewers using a structured, multi-phase process. In the first phase, both reviewers screened the titles and abstracts of all identified articles to assess their relevance based on predefined inclusion and exclusion criteria. Any discrepancies during this initial screening were discussed between the reviewers until consensus was reached; if no agreement could be achieved, a third senior reviewer served as an adjudicator to make the final decision.

2.3. Risk of Bias Assessment

The risk of bias in the included studies was systematically assessed using the QUADAS-2 tool, which is widely utilized to evaluate diagnostic accuracy studies. Four domains were analyzed: Patient Selection (PS), Index Test (IT), Reference Standard (RS), and Flow and Timing (FaT). Each study was evaluated in these domains for both risk of bias and concerns regarding applicability to clinical settings.

Two independent reviewers conducted the assessment for all included studies, applying QUADAS-2 signaling questions to judge the risk of bias in each domain. Responses were categorized as "low risk" when studies reported clear and appropriate methodology, "high risk" when there was evidence of methodological flaws that could affect validity, and "unclear" when reporting was insufficient to permit judgment. Patient selection was evaluated based on the sampling strategy and exclusion criteria; the index test was assessed

according to blinding and predefined thresholds; the reference standard was evaluated for independence and reliability; and the flow and timing were examined in terms of the consistency of testing procedures and time intervals between tests. Disagreements between reviewers were resolved through discussion to reach consensus, and a third senior reviewer was available for adjudication if necessary.

The structured use of QUADAS-2 provided a consistent framework for identifying the strengths and limitations of each study and for interpreting aggregated evidence. This evaluation informed the synthesis of results and supported the validity and reliability of the review's conclusions.

In the second phase, full-text articles deemed potentially eligible were independently reviewed to confirm their inclusion based on detailed methodological criteria. Relevant data—including study objectives, ML techniques used, dataset characteristics, evaluation metrics, and primary findings—were extracted using a standardized data extraction form designed to ensure consistency and reproducibility. Any disagreements related to data extraction or study inclusion were resolved through discussion and, if necessary, adjudicated by a third reviewer.

2.4. Summary of Methodological Rigor

This systematic review was designed to meet high standards of methodological rigor, transparency, and reproducibility. By following the PRISMA 2020 guidelines, implementing a predefined protocol, applying clearly defined eligibility criteria, and conducting multi-reviewer validation at every stage, we ensured that study selection and data extraction were performed consistently and objectively. The use of the QUADAS-2 tool for bias assessment further reinforces the robustness of this review, allowing readers to interpret the presented findings with confidence in their validity.

3. Results

3.1. Risk of Bias Evaluation Results

The majority of studies showed a low risk of bias across most domains. Specifically, 85% of studies exhibited a low risk of bias in patient selection, 93% in the index test, 98% in the reference standard, and 90% in flow and timing. However, a minority of studies raised concerns, particularly in relation to patient selection and procedural consistency within the flow and timing domain. A small number of studies were classified as having an unclear risk due to insufficient methodological reporting.

As summarized in Table 2, these findings indicate that while the overall methodological quality of the included studies was acceptable, residual sources of bias—especially in patient recruitment and test administration—may have influenced specific outcomes. This underscores the importance of cautious interpretation when comparing diagnostic performance across studies.

Table 2. Summary of risk of bias assessment using QUADAS-2.

Domain	Low Risk	High Risk	Unclear Risk
Patient Selection (PS)	85%	10%	5%
Index Test (IT)	93%	0%	7%
Reference Standard (RS)	98%	0%	2%
Flow and Timing (FaT)	90%	5%	5%

Note: Percentages represent the proportion of included studies judged as "low," "high," or "unclear" risk of bias in each QUADAS-2 domain, based on reviewers' consensus assessment.

The QUADAS-2 assessment provides valuable insight into the overall methodological quality of the included studies, revealing that most demonstrated a low risk of bias across key domains. High-quality performance in the Index Test and Reference Standard domains reflects the growing rigor in evaluating ML-based diagnostic models. However, a small subset of studies exhibited high or unclear risk, particularly in the Patient Selection and Flow and Timing domains, underscoring variability in study design, reporting practices, and sample representativeness. These findings highlight the need for standardized reporting frameworks and prospective study designs to ensure the reproducibility and clinical applicability of ML algorithms. By explicitly quantifying and categorizing sources of bias, this review emphasizes both the strengths and limitations of the evidence base, providing a transparent foundation for interpreting diagnostic accuracy results and guiding future research.

3.2. Sensitivity Analyses

Sensitivity analyses were not conducted due to substantial methodological heterogeneity among the included studies, particularly regarding data types, machine learning model architectures, and evaluation protocols. This variability limited the feasibility of subgroup comparisons and prevented the generation of consistent pooled estimates.

3.3. Search Results

A systematic search across the Scopus, ScienceDirect, IEEE Xplore, and Web of Science databases yielded a total of 221 articles. After duplicate removal, 181 unique records remained. Title and abstract screening narrowed this selection to 124 articles explicitly reporting the use of machine learning (ML) for deep vein thrombosis (DVT) detection.

Following the application of predefined eligibility criteria, 32 studies were selected for inclusion in this systematic review. These studies were organized into the following four thematic categories:

- DVT prediction;
- DVT monitoring;
- DVT risk assessment;
- Reference data.

To provide a clear understanding of the research trends and contextual factors influencing machine learning (ML) and deep learning (DL) studies on deep vein thrombosis (DVT), Figures 2–4 provide essential context. Figure 2 illustrates the rapid growth of research activity in this domain, highlighting a marked increase in publications after 2020 which parallels the integration of AI-driven solutions in clinical workflows. This trend underscores the relevance of conducting a timely and comprehensive review, as recent studies have increasingly leveraged DL architectures, multimodal data, and real-time diagnostic tools, reflecting a significant shift in the field's maturity and clinical focus.

Figure 3 highlights the geographic distribution of studies, showing the top ten countries contributing the highest number of publications on ML applications in DVT. This distribution reflects research leadership concentrated in North America and China, driven by the advanced technological infrastructure and healthcare system priorities in these regions. The uneven representation also points to potential dataset and population biases, as many studies originate from single-center or region-specific cohorts. Recognizing this concentration is essential for interpreting model generalizability and identifying opportunities to expand future research into under-represented regions, thereby improving the diversity and applicability of ML-driven clinical solutions.

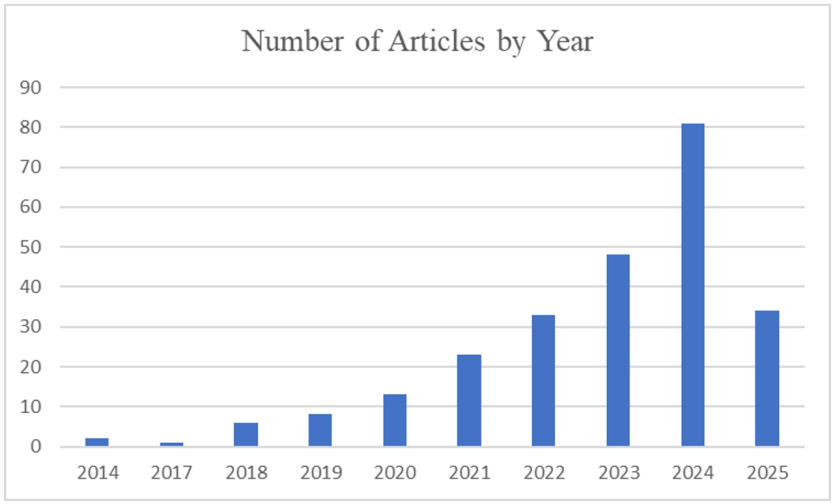

Figure 2. Number of articles by year.

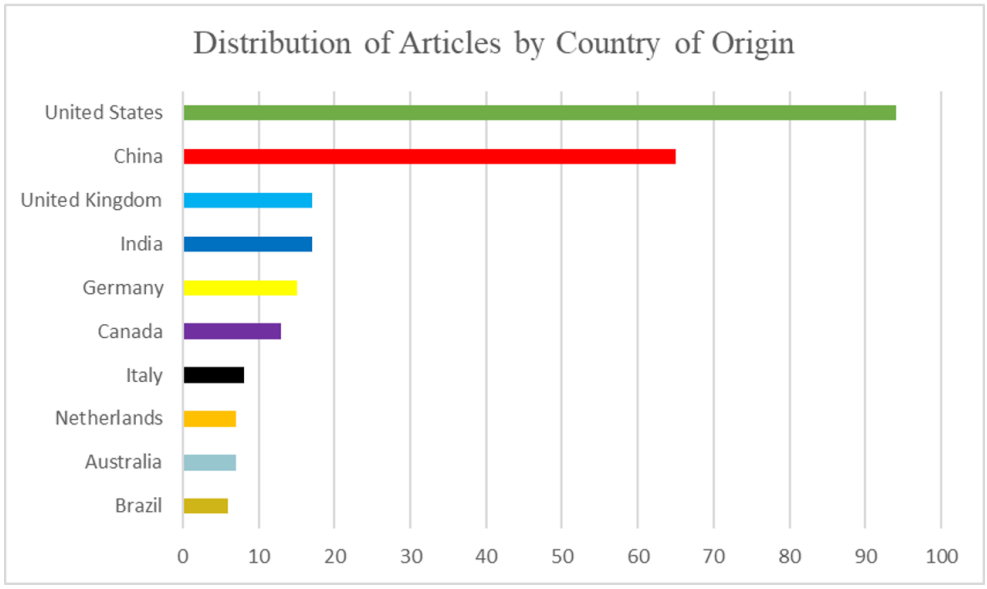

Figure 3. Distribution of articles by country of origin.

Figure 4 presents the thematic categorization of the included studies, organizing them into the four task domains defined in this review: Prediction, monitoring, risk assessment, and reference datasets. This categorization not only summarizes the research focus of the field but also serves as the structural foundation for the Results and Discussion sections, allowing for a clear and systematic synthesis of findings. By framing the literature in these domains, Figure 4 supports a comparative evaluation of methodologies, highlights gaps in validation and clinical translation, and guides readers through the narrative of this review. In addition, Figures 5 and 6 build upon this framework by detailing the algorithm distribution and reported performance metrics, directly addressing the central objective of assessing methodological trends and technical efficacy.

Figure 5 illustrates the relative frequency of machine learning algorithms applied in DVT-related studies, with Random Forest (18%) and Support Vector Machine (15%) emerging as the most frequently adopted approaches. This distribution provides valuable insight into methodological trends and community preferences, reflecting the popularity of tree-based ensemble models for structured clinical data and the historical strength of SVMs in handling small-to-moderate datasets. Understanding this landscape helps to contextualize performance comparisons and underscores the need to explore underuti-

lized algorithms and novel deep learning architectures to enhance generalizability and clinical applicability.

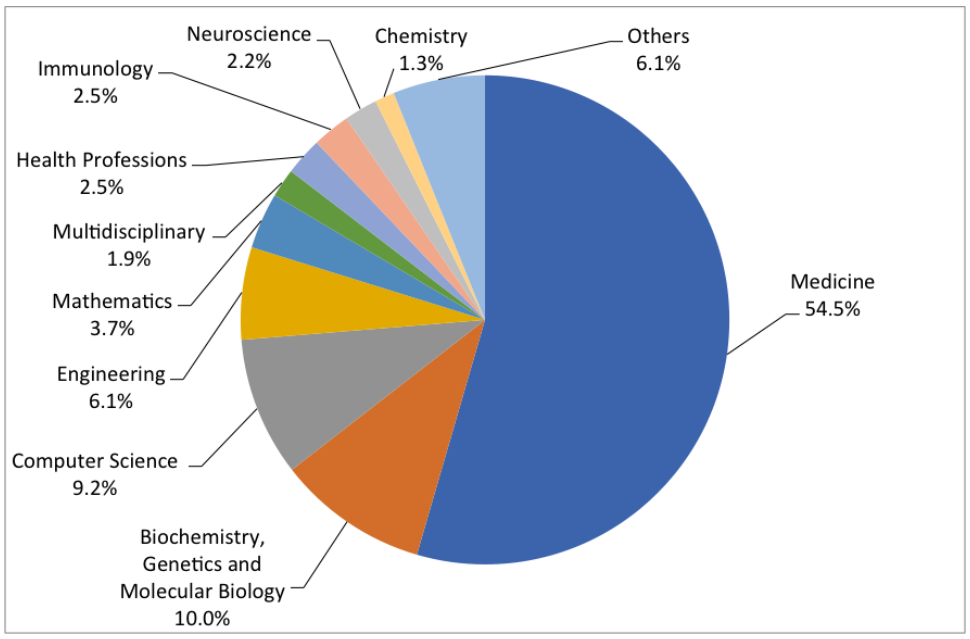

Figure 4. Distribution of articles by subject.

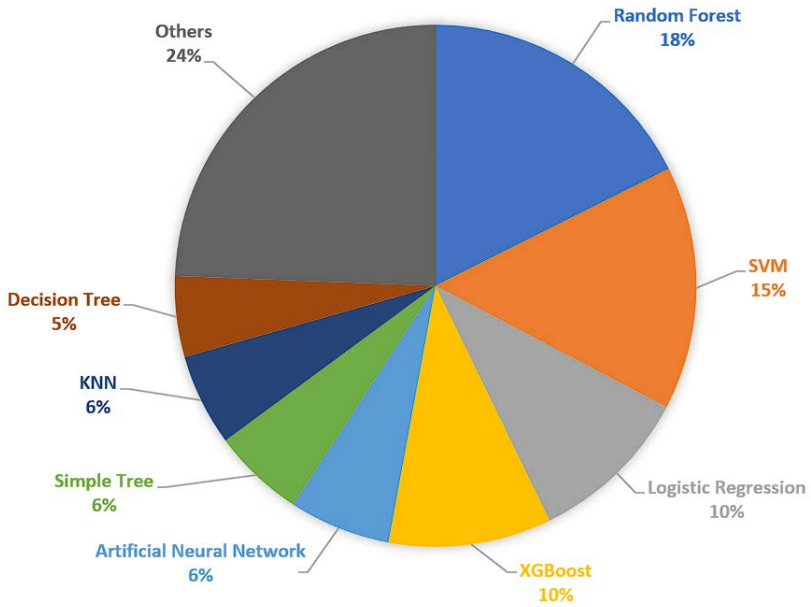

Figure 5. Percentage distribution of ML algorithms used in DVT studies.

Figure 6 summarizes the average accuracy reported for the main machine learning algorithms applied in DVT-related studies. Models based on Random Forest (94%), XGBoost (94%), and Artificial Neural Networks (95%) demonstrated the highest performance metrics reported in the literature. This comparison highlights the strong predictive capability of ensemble tree-based methods and neural architectures, which leverage the feature interactions and non-linear relationships in clinical data. However, it also reveals potential challenges in direct cross-study comparisons, as differences in dataset size, feature selection, and validation strategies can influence the reported accuracy. This figure therefore serves as both a benchmark of existing methodological achievements and a call for standardized

evaluation protocols to ensure fair and clinically meaningful performance assessment across algorithms.

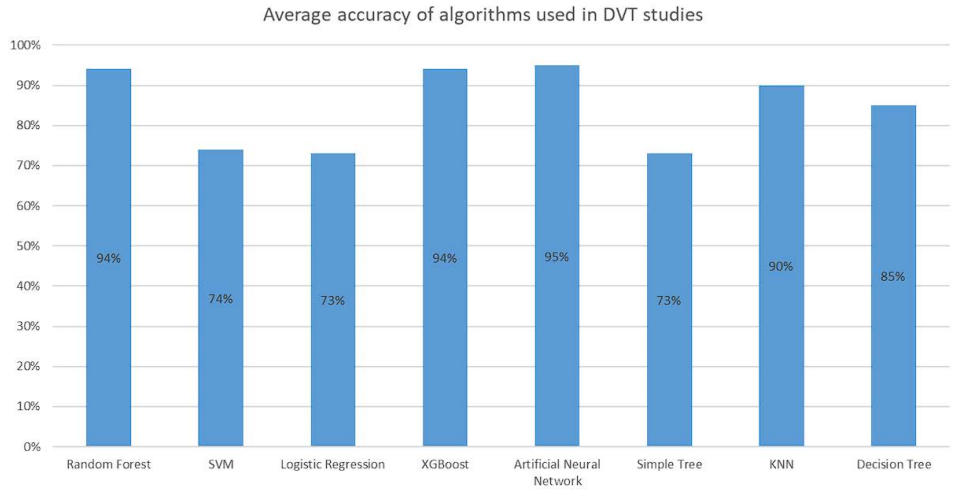

Figure 6. Average accuracy reported in the literature for the main ML algorithms applied in DVT.

Additional visualizations are provided in the Supplementary Materials. These include Figure S1 (journal distribution), Figure S2 (author contributions), Figure S3 (funding sources), Figure S4 (article types), Figure S5 (Co-occurrence Map), and Figure S6 (citation networks), offering a broader bibliographic perspective that supports—but does not distract from—the methodological analysis presented in the main text. There also detailed information regarding journals, authors, and datasets included in this review in the Appendices A–C (Tables A1–A3).

3.4. DVT Prediction

The early and accurate prediction of DVT using ML algorithms has emerged as a prominent research focus in recent years. A recurring pattern observed across the literature is the application of supervised learning models such as Decision Trees (DT), Random Forests (RFs), Support Vector Machines (SVMs), K-Nearest Neighbors (KNN), and simple neural networks [9–20]. These models are typically trained on structured clinical datasets, leveraging features such as patient demographics, vital signs, and laboratory measurements.

A smaller subset of studies has investigated more advanced techniques, including convolutional neural networks (CNNs), ensemble deep learning architectures, and custom hybrid models [9,20–31]. These approaches generally require complex or multimodal inputs—such as medical imaging or real-time physiological signals—and have demonstrated superior performance, particularly in studies with sufficiently large and high-quality datasets.

3.4.1. Algorithms Used

Several studies have emphasized the use of decision trees and logistic regression for analyzing clinical data, incorporating variables such as age, obesity, and prior DVT history. These interpretable and transparent models are often favored in clinical practice due to their simplicity and ease of integration into routine workflows. Most were evaluated through cross-validation and externally validated on independent cohorts, achieving prediction accuracies ranging from 85% to 90%. Notably, one study introduced a clinical decision-support tool, AutoDVT, which combines logistic regression with a user-friendly interface for hospital implementation [10].

Other approaches have explored neural networks trained on medical images, particularly X-rays and ultrasound scans. Convolutional neural networks (CNNs) applied to Doppler and grayscale ultrasound data achieved high diagnostic performance, with reported AUC values of 0.88 and 0.89 and classification accuracies near 75% [11,19,20,23,24,31–33]. One study also demonstrated the successful application of transfer learning, enabling robust predictive performance even on limited datasets [22].

Ensemble models, including Random Forests, Gradient Boosting Machines (GBMs), and K-Nearest Neighbors (KNN), have been shown to perform well on high-dimensional datasets integrating clinical, physiological, and imaging features. However, performance frequently declined in the presence of data imbalance—a recurring challenge in DVT datasets given the relatively low prevalence of positive cases [8,9,34]. Notably, CNN-based models trained with oversampling strategies outperformed traditional classifiers, underscoring the importance of robust data balancing techniques.

To summarize the landscape of ML approaches for DVT prediction, the reviewed studies were categorized by primary model type: Traditional algorithms (e.g., logistic regression, decision trees, SVMs), deep learning models (primarily CNNs), and temporal models (e.g., LSTMs). As shown in Figure 7, traditional ML techniques continue to dominate for structured clinical data, whereas deep learning has gained traction in imaging studies. Temporal models remain underexplored, yet represent a promising avenue for capturing dynamic risk patterns in hospitalized or continuously monitored patients.

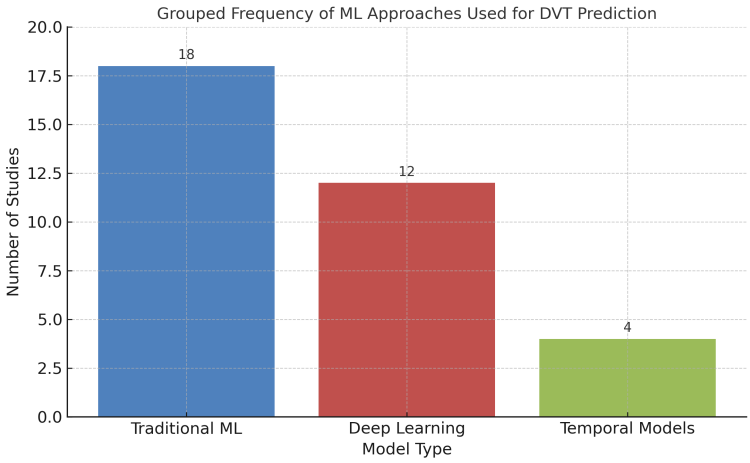

Figure 7. Grouped frequencies of ML model types used for DVT prediction across the reviewed studies.

In summary, conventional ML methods generally achieved accuracies between 80% and 90%, whereas deep learning approaches—particularly those leveraging imaging data—achieved accuracies of up to 95% [31–33]. However, these performance gains were often accompanied by trade-offs in interpretability and computational cost, highlighting the ongoing challenge of balancing model complexity with clinical applicability.

3.4.2. Observed Limitations

Despite promising results, multiple limitations were identified in studies proposing DVT prediction models. Many of these challenges stemmed from reliance on retrospective clinical data, which often contained incomplete, imprecise, or inconsistent information, thereby increasing the risk of bias in ML model development and evaluation [12,16]. Missing values for critical variables—such as anticoagulant treatment history or comorbidities—frequently led to the use of imputation strategies, potentially compromising clinical reliability and reducing model robustness.

Additionally, most algorithms did not incorporate temporal dynamics or contextual factors influencing DVT progression, such as prolonged immobilization, evolving physiological markers, or medication adjustments. The absence of these variables limits the capacity of models to capture clinically meaningful temporal patterns or dynamic trajectories that warrant early intervention.

This gap highlights the importance of adopting temporal modeling techniques—such as recurrent neural networks (RNNs), long short-term memory (LSTM) networks, or temporal convolutional networks (TCNs)—specifically designed to capture time-dependent relationships in clinical data [11,14,29,35]. A small number of studies have applied LSTM networks to longitudinal vital sign records, reporting performance improvements of 5–10 percentage points compared to static models. However, these approaches were limited by the scarcity of labeled time-series datasets and the computational overhead associated with training complex recurrent models in clinical settings.

Furthermore, several studies did not report calibration metrics—such as Brier scores, calibration curves, or reliability diagrams—hindering the evaluation of whether the predicted probabilities accurately reflected observed outcomes. This omission is particularly concerning in DVT applications, where poorly calibrated models may lead to overtreatment (false positives) or missed diagnoses (false negatives), both of which carry significant clinical risk. Table 3 summarizes the best-performing ML approaches for DVT prediction, highlighting the diversity of algorithms, data sources, and clinical applications across studies.

Table 3. Summary of best-performing ML approaches for DVT prediction and their clinical relevance.

Study	Algorithm	Data Type	Objective	Performance	Clinical Insight
[10]	Logistic Regression	Clinical (EHR)	Risk stratification	Accuracy: 87%	AutoDVT tool; interpretable and deployable in hospital triage
[32]	CNN	Doppler Ultrasound	Image-based diagnosis	AUC: 0.89	Enabled point-of-care diagnostics with portable devices
[22]	CNN	Ultrasound + Clinical Metadata	Hybrid imaging + metadata classification	Accuracy: 95%	Used transfer learning; effective with small datasets
[25]	Random Forest	Demographics + Labs	Preoperative risk prediction	Accuracy: 84%	Applied in cancer patient surgical planning
[9]	SVM, KNN, RF	Structured Clinical	General detection	Accuracy: 80–85%	Lower accuracy with imbalanced datasets

3.4.3. Clinical Implications

The integration of ML-based DVT prediction tools into clinical workflows offers several potential advantages that could meaningfully enhance patient care and optimize healthcare system performance:

- Facilitating the early identification of high-risk individuals, particularly in preoperative, emergency, or intensive care settings.
- Optimizing diagnostic resource allocation by prioritizing imaging or laboratory testing for patients with the highest predicted risk.
- Reducing clinician workload through automated, real-time risk stratification integrated into electronic health record (EHR) systems.
- Extending diagnostic support to rural or underserved areas by combining portable imaging devices with on-device ML inference.
- Supporting continuous, in-hospital monitoring by dynamically updating risk assessments as patient conditions evolve.

To fully realize these benefits, future research should prioritize prospective clinical trials, the use of diverse multicenter datasets, and the development of calibrated, interpretable, and patient-centered predictive models designed for seamless integration into existing clinical infrastructures.

Overall, these results illustrate how ML models can be tailored to different clinical contexts, from risk stratification in hospital triage to real-time inpatient monitoring.

3.5. DVT Monitoring

In contrast to diagnostic applications, the use of ML for continuous monitoring of patients at risk for DVT—particularly in postoperative or immobilized contexts—has shown encouraging preliminary results. This potential is especially evident when ML is integrated with wearable devices and mobile health (mHealth) platforms [14,16].

For instance, one study [15] implemented a wearable sensor system connected to a smartphone application, continuously capturing physiological indicators such as heart rate variability, blood pressure trends, and limb movement. This configuration enabled real-time detection of anomalies associated with thrombotic risk, facilitating timely clinical intervention. Clinically, the potential of this approach suggests that patients recovering from orthopedic surgery or with known coagulation disorders could benefit from remote, non-invasive monitoring, potentially reducing hospital stays and supporting proactive outpatient management.

3.5.1. Real-Time Approaches

Several studies have explored unsupervised learning techniques to identify early physiological anomalies that may precede thrombotic events. For example, clustering algorithms such as k-means and hierarchical clustering have been used to group patients with similar risk profiles or to detect outlier patterns which are indicative of early DVT onset [27,29,30]. These models have shown particular value in postoperative monitoring, where sudden changes in cardiovascular dynamics can signal clot formation. Principal Component Analysis (PCA) has also been employed in several studies for dimensionality reduction and to isolate the most informative features from noisy real-time data streams [15,21,36]. This technique allows clinicians to focus on key physiological indicators—such as peripheral temperature drops or sharp decreases in blood flow velocity—that are strongly associated with venous obstruction.

Although less frequently applied in this context, deep learning models have demonstrated significant promise for real-time biomedical signal interpretation. For instance, a convolutional neural network (CNN) was applied to Doppler ultrasound signals acquired via wearable probes, achieving an accuracy of 89% in detecting venous obstructions in the lower extremities [22,31]. These findings suggest that ambulatory Doppler monitoring could evolve into a practical clinical tool, particularly for high-risk outpatient populations including cancer patients and individuals with a history of DVT.

Notably, some models have demonstrated the ability to detect thrombotic patterns hours before the onset of clinical symptoms, underscoring their potential value in preventive medicine. Such capabilities could enable earlier initiation of anticoagulant therapy or prompt confirmatory imaging, thereby improving patient outcomes [23,24,37].

Figure 8 illustrates the desired data flow for ML-based DVT monitoring. Sensors capture physiological signals, which are preprocessed and subjected to PCA to isolate the most relevant features. These features are subsequently analyzed by clustering algorithms to assess thrombotic risk. Finally, model outputs can trigger real-time alerts that inform timely clinical decision making, including the initiation of preventive measures.

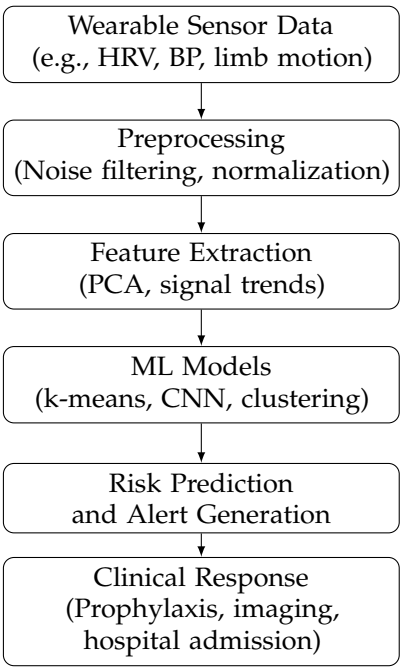

Figure 8. End-to-end ML-based DVT monitoring workflow.

3.5.2. Limitations of Automated Monitoring

Despite recent advances, automated DVT monitoring approaches still face several critical limitations that hinder their widespread clinical adoption. A recurrent challenge is the low specificity of many wearable sensor systems. For example, one study reported that over 30% of alerts generated by an ML model were false positives, frequently triggered by non-pathological physiological fluctuations such as physical activity or benign arrhythmias [15]. Such frequent false alarms may contribute to alert fatigue among both patients and healthcare providers, undermining trust in the reliability of these monitoring systems.

Economic feasibility represents another significant barrier to large-scale implementation. High upfront costs related to device development, calibration, and distribution—combined with ongoing maintenance requirements for wearable-based ML systems, particularly those incorporating imaging or Doppler technologies—may limit adoption in resource-constrained healthcare environments [33,38].

Moreover, substantial inter-individual variability in physiological signals—driven by factors such as age, sex, comorbidities, medications, hydration status, and body posture—often results in models with limited generalizability across populations. Several studies have reported that thresholds which were effective for thrombotic risk detection in one cohort produced inconsistent or misleading results in others, such as when comparing younger athletic populations with elderly, immobilized patients.

To address these challenges, researchers have proposed personalized ML models trained on individualized datasets or demographically similar cohorts. Such models can incorporate adaptive learning algorithms that dynamically adjust thresholds based on a patient's baseline physiological profile. Clinically, this approach has the potential to deliver tailored monitoring strategies that align with each patient's unique risk factors. However, this personalization also introduces ethical and logistical challenges related to the use, storage, and transmission of sensitive health data, particularly in cloud-based systems for real-time analytics [17,18,39]. Table 4 summarizes the most relevant ML-based DVT monitoring approaches, detailing the range of data sources—spanning wearable sensors, Doppler ultrasound, multimodal ICU data, and electronic health records—and reflecting the diversity of clinical contexts in which continuous monitoring is applied.

Table 4. Summary of best ML approaches for DVT monitoring in the reviewed literature and their clinical implications.

Study	ML Method	Data Source	Clinical Context	Performance	Clinical Insights
[15]	Clustering (Unsupervised)	Wearable sensors (HRV, BP)	Postoperative recovery	Not reported	Enabled real-time alert generation via smartphone.
[22]	CNN	Doppler ultrasound (wearable)	Outpatient oncology follow-up	Accuracy: 89%	Achieved pre-symptomatic detection of venous obstruction.
[21]	PCA + Clustering	Multimodal (physiological + clinical)	ICU post-surgery	N/A	Improved interpretability of complex signal inputs.
[30]	Hierarchical Clustering	EHR + vitals	In-hospital surveillance	N/A	Identified subgroups with elevated thrombotic risk.
[24]	CNN	Doppler flow recordings	Telemedicine DVT programs	AUC: 0.87	Integration into remote triage systems for at-risk populations.

3.5.3. Clinical Implications

The integration of ML-based monitoring tools into routine clinical workflows offers substantial potential to improve patient outcomes and optimize healthcare resource utilization:

- Facilitating early detection and intervention for high-risk patients during outpatient care or home-based recovery.
- Reducing hospital readmissions, particularly after orthopedic surgeries or cancer treatments.
- Enhancing diagnostic efficiency through ML-driven alerts that prioritize confirmatory testing and imaging resources.
- Enabling the design of closed-loop systems capable of autonomously recommending—or even initiating—prophylactic interventions based on real-time risk assessment.

To achieve these benefits, successful implementations will require rigorous validation through multicenter clinical trials, close collaboration between clinicians and data scientists, and strict adherence to regulatory frameworks to ensure safety, reliability, and ethical transparency in the deployment of these models.

Overall, these findings highlight the clinical potential of ML for continuous DVT monitoring. While wearable and telemedicine-based approaches show promise for outpatient care and early detection, unsupervised clustering methods offer insights for patient stratification and risk identification within hospital and ICU settings.

3.6. DVT Risk Calculation

An important area of research focuses on estimating the risk of DVT using predictive ML models trained on historical clinical data and validated risk factors. Unlike diagnostic models, which aim to identify existing DVT cases, these predictive systems estimate the likelihood of thrombus development over time.

The primary objective of these models is to stratify patients by risk level, enabling early and targeted interventions that may prevent thrombosis and reduce the likelihood of severe complications. Numerous studies have demonstrated that these models enhance clinical decision making by integrating a wide range of variables, including demographic data, comorbidities, and laboratory test results [1,4,25,29]. Clinically, this stratification supports the selective administration of prophylactic interventions—such as anticoagulation, compression therapy, or perioperative adjustments—to patients with elevated predicted risk, thereby optimizing resource allocation and minimizing overtreatment.

3.6.1. Risk Stratification Models

Unlike diagnostic algorithms that rely on imaging or acute symptoms, risk stratification models often employ logistic regression and Bayesian networks to provide a more precise assessment of individual patient risk [32,40,41]. These models typically process

variables such as age, sex, body mass index (BMI), cancer history, duration of immobilization, and family medical history, classifying patients into low-, moderate-, or high-risk categories. The reported accuracy rates for these methods are generally above 81% [27,40].

Bayesian network-based approaches have shown particular promise in addressing the probabilistic and uncertain nature of clinical data. By modeling interactions among multiple risk factors—such as oral contraceptive use, genetic predisposition, and physical trauma—these models produce nuanced and individualized risk estimates. This capability makes them especially effective for identifying borderline or atypical patient profiles that may not be flagged by traditional rule-based tools, such as the Wells score [17,18,42].

Recent advances also include the application of ensemble methods, such as Random Forests and Gradient Boosting Machines (GBMs), to predict the risk of DVT in high-risk populations such as oncology patients, orthopedic surgery recipients, and ICU patients. These models incorporate dynamic clinical parameters—including intraoperative blood loss, post-surgical mobility, and inflammatory markers—to deliver more accurate and context-specific predictions, often achieving accuracy levels of up to 87% [17,18,27,31,32,37]. When integrated into electronic health record (EHR) systems, these algorithms can generate automated alerts and clinical decision-support tools to guide thromboprophylaxis in real-time.

To illustrate the influence of individual clinical features on ML-based risk stratification, Figure 9 presents an aggregated estimate of feature importance values synthesized from multiple studies, rather than a single dataset [17,27,32,40,42]. Variables such as cancer status, immobility, and age consistently emerge as top contributors, particularly in Random Forest and GBM models. While this figure is illustrative, it reflects consensus trends in feature relevance reported across the literature.

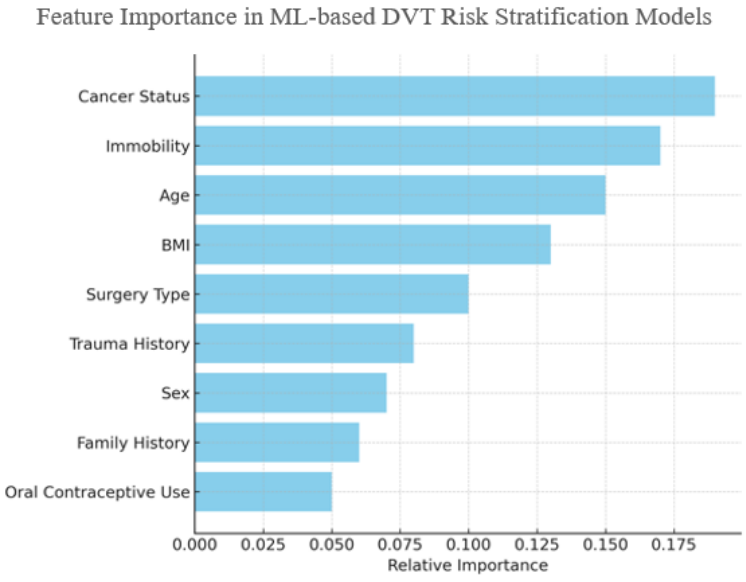

Figure 9. Relative importance of clinical features in machine learning models for DVT risk stratification (feature importance values were approximated based on frequency and prominence in the reviewed literature, not on a unified dataset).

3.6.2. Limitations of Risk Models

Despite their potential, existing models for DVT risk stratification face several important limitations. A common challenge is the reliance on patient cohorts derived from a single hospital or geographically restricted datasets, which limits both generalizability and external validity [17,27]. This limitation becomes particularly critical when deploying models across diverse healthcare environments, where patient demographics, clinical practices, and diagnostic standards can vary substantially.

Another major limitation is the insufficient integration of genomic, proteomic, and molecular biomarkers, which play crucial roles in individual susceptibility to DVT. The exclusion of these variables constrains predictive accuracy and may overlook essential biological contributors [17,26,39]. For example, genetic variants such as factor V Leiden, prothrombin G20210A, and deficiencies in natural anticoagulants (e.g., antithrombin III, protein C, protein S) are well-established risk factors, yet are often absent from prediction pipelines due to limited data availability.

Furthermore, while many studies have reported favorable performance metrics—such as accuracy or AUC—there is a consistent lack of external validation using prospective cohorts, and very few models have been tested within real-world clinical workflows. To ensure clinical applicability, future models must undergo rigorous prospective validation, be integrated seamlessly into hospital information systems, and incorporate intuitive, user-centered interfaces that facilitate clinical decision making. Only under such conditions can these predictive tools progress from academic research to reliable instruments in preventive vascular medicine.

Table 5 provides an overview of the most relevant ML-based models developed for DVT risk stratification. These studies primarily focused on patient-specific risk factors—including demographics, comorbidities, lifestyle, and perioperative conditions—to enable the early identification and classification of individuals at increased risk.

Table 5. Summary of best ML-based DVT risk stratification models in the reviewed literature and their clinical applications.

Study	ML Model	Key Variables Used	Patient Population	Performance	Clinical Application
[40]	Logistic Regression	Age, BMI, prior DVT, immobility	Surgical inpatients	Accuracy: 81%	Risk group classification for perioperative prophylaxis
[17]	Bayesian Network	Cancer status, contraceptive use, trauma history	Oncology cohort	Not reported	Uncertainty modeling in complex patient cases
[32]	Random Forest	Labs, vitals, comorbidities	ICU patients	AUC: 0.87	Trigger alerts for intensive thromboprophylaxis
[27]	GBM	Surgery type, mobility, age, sex	Ortho/trauma patients	Accuracy: 85%	Supports stratification in fast-track discharge protocols
[42]	Bayesian Network	Demographics, family history, lifestyle	General population	Not available	Screening tool for outpatient DVT prevention plans

Overall, the studies highlight the adaptability of ML approaches to diverse patient populations, ranging from general outpatients to high-risk surgical and oncology cohorts. Logistic regression and tree-based methods offer robust performance in structured hospital data, while Bayesian networks provide additional interpretability in cases involving uncertainty and complex variable interactions. Together, these approaches underscore the roles of ML in guiding targeted thromboprophylaxis and personalized prevention strategies.

3.7. Reference Data

A consistent finding across all reviewed studies is that the performance and reliability of ML models are highly dependent on the quality and quantity of the reference datasets used for training and validation. This section underscores the critical role of datasets in shaping model development, predictive accuracy, and clinical applicability. The volume, heterogeneity, and provenance of training data directly determine a model's capacity to generalize across diverse patient populations and clinical settings. Therefore, both dataset size and diversity represent essential determinants of a model's robustness and generalizability [29,30,39].

3.7.1. Data Sources

The reviewed studies predominantly relied on clinical datasets sourced from local clinics [11–15,29,30,32–34,43] and hospital databases [17,18,21–28,31], with a limited number of studies employing synthetically generated datasets to address data scarcity concerns [9,10]. Frequently referenced sources included the UK-VTE registry and institutional records from German hospitals [17,18,21–28]. These datasets typically included patient demographics, medical histories, treatment regimens, laboratory test results, and routinely collected physiological parameters.

Several studies have reported improved predictive performance through the integration of multimodal data sources. Some approaches combined imaging modalities—such as Doppler ultrasound—with physiological data from wearable devices and structured clinical records, resulting in enhanced model accuracy and interpretability [29,32,33]. For instance, one study demonstrated that incorporating Doppler imaging with clinical variables increased predictive accuracy by approximately 15% [36].

3.7.2. Data Limitations

One of the most critical challenges identified across the reviewed studies is the lack of dataset standardization and variability in data quality. Heterogeneity across sources introduces inconsistencies that hinder the generalizability of ML models. In numerous cases, electronic health records (EHRs) contained missing values, entry errors, or inconsistently defined variables, all of which compromised model reliability. Additionally, several studies reported insufficient preprocessing and data cleaning procedures, potentially amplifying noise or introducing bias.

Furthermore, most studies did not provide detailed information regarding the provenance of datasets, raising concerns about the reliability of predictions, monitoring strategies, or risk estimation models. Limited transparency in data origin also complicates external validation and poses challenges for regulatory approval. The scarcity of open-access, well-annotated benchmark datasets further restricts reproducibility and limits the ability to conduct comparative evaluations of algorithms. This lack of data-sharing initiatives prevents the research community from establishing standardized baselines, impeding progress toward robust and clinically applicable ML solutions.

These findings emphasize the urgent need for standardized datasets, rigorous documentation protocols, and collaborative data-sharing frameworks to strengthen model development, evaluation, and clinical translation. Table 6 summarizes the relationships between authors and the ML algorithms analyzed in this review.

Table 6. Relationships between the authors and ML algorithms used in DVT studies.

Author	Model–Algorithm
[1]	Artificial Neural Network
[2]	XGBoost, Gradient Boosting, Decision Tree
[3]	SVM
[4]	Simple Tree, KNN, Random Forest, SVM
[5]	Artificial Neural Network
[6]	SVM, Simple Tree
[7]	Simple Tree, KNN, Random Forest, SVM
[8]	Simple Tree, Fine KNN, Random Forest, SVM, Artificial Neural Network
[9]	Artificial Neural Network, SVM, Decision Tree, KNN, Random Forest, XGBoost, Gradient Boosting

Table 6. *Cont.*

Author	Model–Algorithm
[10]	SVM, Decision Trees, Extra Trees, Random Forest, KNN
[11]	SVM, Decision Trees, MLP-NN
[12]	ResNet-NN
[13]	Logistic Regression, XGBoost
[14]	ResNet-NN
[15]	Random Forest, Linear Regression, Artificial Neural Network
[16]	XGBoost, Logistic Regression, Random Forest, SVM
[17]	XGBoost, Random Forest, Logistic Regression
[18]	Convolutional Neural Network
[19]	Proposed Neural Network
[20]	Convolutional Neural Network, Logistic Regression, SVM, Random Forest, XGBoost
[21]	Decision Tree, Random Forest
[22]	Random Forest, XGBoost, SVM, Extra Trees
[23]	Simple Tree, KNN, Random Forest, SVM
[24]	Convolutional Neural Network
[25]	SVM, Simple Tree
[26]	Random Forest
[27]	SVM
[28]	Simple Tree, KNN, Random Forest, SVM
[29]	Convolutional Neural Network
[30]	XGBoost, Gradient Boosting, Decision Tree
[31]	Convolutional Neural Network, Recurrent Neural Network
[32]	Simple Tree, KNN, Random Forest, SVM
[33]	Artificial Neural Network
[34]	Artificial Neural Network
[35]	XGBoost, Random Forest, SVM, Logistic Regression
[36]	Logistic Regression
[37]	Random Forest, Logistic Regression, Gradient Boosting, Decision Tree, XGBoost, KNN
[38]	Logistic Regression, Random Forest, SVM, XGBoost, Naive Bayes
[39]	Hybrid Model
[40]	Random Forest, XGBoost, SVM, Gradient Boosting, Decision Tree, Logistic Regression
[41]	Logistic Regression, SVM, Random Forest, Artificial Neural Network
[42]	XGBoost, Logistic Regression, Random Forest, SVM
[43]	ResNet-NN
[44]	Convolutional Neural Network, Recurrent Neural Network
[45]	Simple Tree, KNN, Random Forest, SVM, Artificial Neural Network
[46]	Graph Attention Networks
[47]	VGG16, VGG19, ResNet, GPT
[48]	Unspecified
[49]	Unspecified
[50]	SHAP, LIME, Decision Tree, Logistic Regression
[51]	Black-Box Models
[52]	GAN, RGAN, RCGAN
[53]	Digital Twins Model
[54]	Unspecified
[55]	Random Forest, XGBoost, SVM, Naive Bayes
[56]	Proposed Model
[57]	Unspecified
[58]	Random Forest, SVM, Gradient Boosting
[59]	Unspecified
[60]	Logistic Regression, Decision Tree, Random Forest

Table 6. *Cont.*

Author	Model–Algorithm
[61]	Proposed Neural Network
[62]	Logistic Regression, Random Forest, Artificial Neural Network, XGBoost
[63]	Logistic Regression, Random Forest, Gradient Boosting
[64]	Logistic Regression
[65]	Logistic Regression, Random Forest, SVM, XGBoost
[66]	Convolutional Neural Network

3.8. Reporting Biases

Potential reporting biases were carefully considered during the review process. The majority of the included studies reported positive findings regarding the application of ML models for the early detection, risk prediction, and monitoring of DVT. This trend raises concerns about potential publication bias, as studies reporting negative or non-significant results may be under-represented in the literature.

Furthermore, reliance on bibliographic databases that primarily index peer-reviewed journals may have contributed to the exclusion of unpublished or gray literature, potentially overestimating the effectiveness and robustness of evaluated models. To address this limitation, future systematic reviews should explicitly incorporate preprints, conference proceedings, and institutional or technical reports to mitigate reporting bias and provide a more comprehensive and balanced assessment of ML-based approaches in this domain.

3.9. Certainty of Evidence

The overall certainty of the evidence synthesized in this review was assessed to be moderate. Although numerous studies demonstrated the promising diagnostic performance of ML models for DVT detection and risk prediction, several factors limit confidence in these findings. Key limitations include methodological heterogeneity, reliance on retrospective datasets, insufficient external validation, and unclear risk-of-bias assessments in some studies, particularly regarding patient selection, study flow, and timing.

Furthermore, the absence of standardized reporting frameworks and limited transparency in model development pipelines challenge reproducibility and hinder clinical translation. Future prospective multicenter investigations adhering to standardized protocols and transparent reporting practices are essential to strengthen the evidence base and support the safe and effective integration of ML models into DVT management.

4. Discussion

This review provides a critical synthesis of ML applications for the prediction, monitoring, and risk stratification of DVT. The findings underscore the growing use of supervised learning algorithms, including Decision Trees, Support Vector Machines, and ensemble methods, alongside more recent investigations into deep learning (DL) architectures for both diagnostic and patient monitoring purposes.

Although ML models consistently demonstrated promising performance—particularly DL-based approaches, with reported accuracies of up to 95%—several limitations undermine their clinical readiness. Traditional models leveraging structured clinical variables achieved robust but comparatively lower accuracy, while convolutional neural networks (CNNs) and other complex architectures offered superior pattern recognition in imaging and physiological data at the cost of increased computational complexity, decreased interpretability, and substantial data requirements.

Compared with ML applications in other cardiovascular conditions, such as stroke or myocardial infarction, ML research in DVT remains less mature and fragmented. Many

studies relied on small single-center datasets, lacked external validation, and reported inconsistencies in data quality and preprocessing pipelines. These limitations raise concerns about model overfitting, particularly when high performance metrics are presented without independent validation.

A key observation is the limited incorporation of temporal information into predictive models. Despite the clinical significance of time-dependent risk factors—such as immobilization duration or changes in pharmacological management—most models analyzed patient data as static snapshots, overlooking dynamic disease trajectories. While wearable technologies and continuous monitoring approaches show promise, they remain challenged by high false positive rates and inter-patient variability in physiological signals.

Risk stratification models demonstrated utility in identifying high-risk patients and guiding early preventive interventions. Logistic regression and Bayesian network-based approaches were particularly effective in managing uncertainty and probabilistic relationships among clinical variables. However, the exclusion of genomic, proteomic, and other molecular biomarkers was a recurring limitation, reducing the ability of current models to capture the full heterogeneity of thrombotic risk.

Data limitations were pervasive across studies. The majority of models were trained on retrospective clinical records, often containing missing values or incomplete information. The lack of standardized, multicenter, and open-access datasets hinders reproducibility, comparative benchmarking, and external validation efforts. Although some studies have explored synthetic or multimodal data integration to overcome these barriers, such methods risk introducing biases that further challenge model reliability and real-world applicability.

From a clinical perspective, ML systems have substantial potential to enhance diagnostic workflows by improving early detection rates, extending advanced diagnostic capabilities to resource-limited settings, and enabling continuous monitoring of at-risk populations. However, translating this potential into practice requires standardized data collection, robust external validation, and the incorporation of temporal and biological complexity to ensure clinical safety and trustworthiness.

To accelerate clinical adoption, future research should prioritize the development of interoperable datasets, the inclusion of multi-omics data, and the integration of interpretable ML models into existing clinical workflows. Ethical considerations—such as data privacy, explainability, and regulatory compliance—must also be addressed to ensure safe and equitable implementation. Additionally, interdisciplinary collaboration between clinicians, data scientists, and regulatory bodies will be key to achieving clinically validated, scalable ML solutions for DVT management.

In conclusion, while ML has transformative potential for the early detection, prevention, and monitoring of DVT, significant technical, clinical, and ethical challenges remain. Addressing these challenges is critical to progressing from proof-of-concept models to validated, deployable decision-support tools which are capable of improving patient outcomes in real-world healthcare environments.

5. Trends and Future Work

A major emerging trend in DVT research is the integration of multimodal datasets to enhance the accuracy, robustness, and generalizability of ML models. While earlier studies primarily relied on single-source data—such as Doppler ultrasound images or structured electronic health records—recent investigations have increasingly demonstrated that combining heterogeneous inputs—including genetic information, wearable sensor outputs, laboratory results, and imaging modalities—substantially improves predictive performance [9,21,27]. This shift is enabled by the availability of high-dimensional clin-

ical datasets and the development of advanced ML architectures capable of processing heterogeneous data streams.

Advanced deep learning techniques, such as transformers and graph neural networks (GNNs), have shown strong potential for modeling nonlinear relationships and spatiotemporal dependencies, making them well-suited for multimodal data integration in clinical contexts [32,34,46]. Future research should focus on building scalable multimodal ML frameworks which are capable of merging disparate data sources while addressing challenges related to data harmonization, interoperability, and standardization [13–15]. Innovations in federated feature engineering, semantic ontologies, and automated data harmonization pipelines will be key enablers of this paradigm [11,12].

Another rapidly growing area is the development of adaptive and personalized ML models tailored to individual patient characteristics. Evidence indicates that inter-patient variability in genetic, physiological, and environmental factors reduces the effectiveness of generalized prediction models [10,15,33]. Consequently, transfer learning, meta-learning, and few-shot learning strategies are gaining traction as approaches to enable dynamic model adaptation with minimal retraining [12,47,48]. In the near future, lifelong learning systems integrated with edge–cloud infrastructures could enable continuous model refinement, ensuring that predictions remain aligned with patient-specific risk profiles in real-time.

Federated learning is another transformative research direction, allowing for decentralized model training across multiple institutions without centralized data pooling. This approach preserves patient privacy, reduces data governance barriers, and supports greater model generalizability across diverse populations [11,14,35]. As federated networks scale globally, they hold potential to form collaborative AI ecosystems that are capable of adapting dynamically to epidemiological trends and regional variations in DVT prevalence [18,27,29,31,49].

Explainable Artificial Intelligence (XAI) has emerged as a critical priority for clinical ML deployment. The opacity of complex models continues to undermine clinician trust and presents regulatory hurdles [15,16]. Cutting-edge XAI methods, such as attention-based visualization, counterfactual reasoning, and causal inference, are essential to make model outputs interpretable and clinically actionable [50,51]. Future ML frameworks must embed interpretability at their core to meet clinical decision-making standards and regulatory requirements.

Synthetic data generation represents another promising avenue to address data scarcity and bias in DVT research. Techniques such as Generative Adversarial Networks (GANs) and diffusion models are being explored for the production of realistic, privacy-preserving datasets [44,52]. However, their widespread adoption will require rigorous validation frameworks, standardized benchmarking protocols, and regulatory acceptance [32,38,39]. Future efforts should focus on evaluating synthetic datasets against prospective clinical outcomes to ensure reliability and real-world utility.

Looking ahead, the integration of ML models with wearable biosensors, remote monitoring systems, and digital twin platforms could enable predictive, preventive, and personalized medicine in the context of DVT care [53,54]. Achieving this vision will require technological maturity, ethical safeguards, robust governance frameworks, and close interdisciplinary collaboration between clinicians, data scientists, biomedical engineers, and policymakers. These efforts will be critical for translating technical innovation into tangible clinical impact.

6. Limitations and Clinical Implications

6.1. Data Access and Quality

A primary limitation of this review is its reliance on literature indexed in a restricted set of scientific databases that meet specific criteria for access, indexing, and availability. Consequently, studies published in languages other than English or in regional, non-indexed, or domain-specific journals may have been unintentionally excluded. This limitation introduces potential selection bias and restricts the cultural and geographical diversity of insights, particularly concerning ML applications in under-resourced healthcare environments.

The reviewed studies also displayed substantial methodological heterogeneity, particularly regarding ML model development, validation, and reporting practices. Differences in dataset composition, preprocessing pipelines, hyperparameter tuning, and evaluation metrics (e.g., accuracy vs. AUC) complicated direct cross-study comparisons. Variability in feature engineering approaches, model selection (e.g., logistic regression, SVMs, deep neural networks), and validation strategies further hindered efforts to synthesize the findings, establish benchmarking standards, or identify clearly superior methodologies. This methodological inconsistency represents a key barrier to reproducibility and the consolidation of evidence.

Another limitation stems from the widespread use of retrospective datasets derived from electronic health records (EHRs), which are frequently affected by missing values, inconsistent data structures, and variability in data quality. In many studies, data preprocessing steps—including imputation methods, outlier management, or normalization techniques—were insufficiently documented or inconsistently applied, reducing reproducibility and model robustness. The absence of prospective validation in most studies further complicates the assessment of whether reported performance metrics would translate to real-world clinical workflows, where patient variables are dynamic and often poorly standardized. These limitations underscore the urgent need for future research to adopt prospective multicenter validation protocols, harmonized preprocessing pipelines, and comprehensive data quality assessments to strengthen the reliability and generalizability of ML-based solutions in DVT management.

6.2. Model Evaluation and Generalizability

A major limitation identified across the reviewed studies is the lack of rigorous external validation. Only a small subset of investigations evaluated their models on datasets from different institutions, geographic regions, or patient populations. The absence of external validation undermines confidence in model generalizability, as variations in population health characteristics, diagnostic protocols, and data acquisition techniques can significantly impact performance. Without independent validation, it remains uncertain whether these models can maintain predictive accuracy across diverse healthcare environments, limiting their readiness for clinical deployment.

Additionally, the studies employed heterogeneous performance metrics—including accuracy, sensitivity, specificity, and AUC—without clear consensus on which measures most effectively capture clinical utility for DVT prediction or diagnosis. In many cases, essential elements such as calibration metrics, decision thresholds, or strategies for managing class imbalance were under-reported, reducing interpretability and clinical relevance. This inconsistency complicates cross-study comparisons and risks generating overly optimistic interpretations of model performance.

To address these gaps, future research should prioritize the development of standardized evaluation protocols and benchmarking frameworks. Establishing shared datasets, harmonized metrics, and transparent reporting guidelines would enhance reproducibility

and comparability across studies. Emphasizing external validation and independent testing will be essential to demonstrate the robustness and reliability of ML-based approaches, ultimately accelerating their translation into routine DVT care.

6.3. Clinical Implications

The integration of ML models into clinical workflows for DVT management holds transformative potential across diagnostic, predictive, and monitoring domains. This review demonstrates how ML approaches—from traditional classifiers such as logistic regression and Decision Trees to advanced architectures like Convolutional Neural Networks (CNNs) and ensemble methods—can complement existing diagnostic pathways, enhance efficiency, and support personalized care delivery.

(a) Early Detection:

ML-based diagnostic systems trained on imaging modalities such as Doppler ultrasound or compression sonography offer promising alternatives to conventional diagnostic workflows. These algorithms can automate image interpretation, highlight thrombotic regions, and provide diagnostic support in settings with limited radiological expertise. By enabling earlier and more accessible detection in primary care or resource-constrained environments, these systems have the potential to reduce diagnostic delays and improve patient outcomes.

(b) Risk Stratification:

Predictive ML models that incorporate electronic health records, demographic data, and comorbidities have demonstrated strong potential for identifying individuals at elevated risk of DVT. Such stratification tools could guide prophylactic care, reduce unnecessary interventions, and optimize resource allocation. When implemented as triage solutions, these models can inform preventive strategies, including targeted anticoagulation and compression therapies, improving both safety and efficiency in clinical decision making.

(c) Real-Time Monitoring:

Although still in early development, wearable biosensor technologies integrated with ML algorithms represent a promising strategy for continuous patient monitoring. These systems analyze real-time physiological signals—such as blood flow velocity, limb movement, and heart rate variability—to detect thrombotic risk patterns before the onset of symptoms. By generating actionable alerts, these systems could enable proactive intervention and potentially prevent thrombotic events in high-risk populations.

Despite these opportunities, several barriers must be overcome to achieve large-scale clinical adoption:

- Data Quality and Standardization: Current models are often trained on retrospective, heterogeneous datasets, reducing their generalizability. The lack of standardized protocols for data collection and preprocessing undermines reproducibility and scalability across healthcare settings.
- Model Interpretability: The “black-box” nature of complex ML models limits clinical trust, particularly in high-stakes diagnostic scenarios. Incorporating Explainable AI (XAI) methods is essential to improve interpretability, regulatory compliance, and clinical acceptance.
- Regulatory and Ethical Considerations: Challenges persist regarding data privacy, model transparency, and accountability in decision making. Developing robust regulatory guidelines, ethical frameworks, and governance models is crucial to ensure the responsible implementation of ML in patient care.
- Validation and Clinical Translation: Most reviewed models lack prospective validation or large-scale clinical trials, limiting their applicability in real-world settings. Future re-

search should prioritize rigorous multicenter evaluations and real-world deployment studies to confirm their generalizability and safety.

6.4. Real-World Usage

Although ML approaches for DVT detection, monitoring, and risk stratification have demonstrated encouraging results, their real-world clinical adoption remains limited. Most studies analyzed in this review relied on retrospective single-center datasets or geographically restricted cohorts, raising concerns about model generalizability across diverse patient populations and healthcare environments. While these systems offer potential advantages—such as automated risk assessment, rapid interpretation of ultrasound imaging, and continuous monitoring via wearable sensors—several challenges impede their translation into routine practice.

A key barrier is the inability of many ML models to replicate their reported performance outside the original study context, largely due to heterogeneity in patient demographics, imaging devices, and clinical protocols. Additionally, integration with existing electronic health record (EHR) infrastructures and compliance with data privacy regulations pose significant implementation challenges. The limited interpretability of complex models further undermines clinical trust, particularly in the context of high-stakes decisions such as anticoagulation therapy planning.

Another critical limitation is the lack of prospective validation and randomized controlled trials, which are essential to verify clinical safety and performance. Figure 10 illustrates a representative workflow for ML-enabled DVT care and highlights common barriers to deployment. To address these issues, future research should prioritize multicenter collaborations, the development of large heterogeneous datasets, and rigorous external validation. Emerging methodologies, including federated learning and explainable AI (XAI), offer promising avenues to overcome data-sharing constraints, improve transparency, and enhance clinician confidence, ultimately facilitating the integration of ML-driven tools into real-world DVT management.

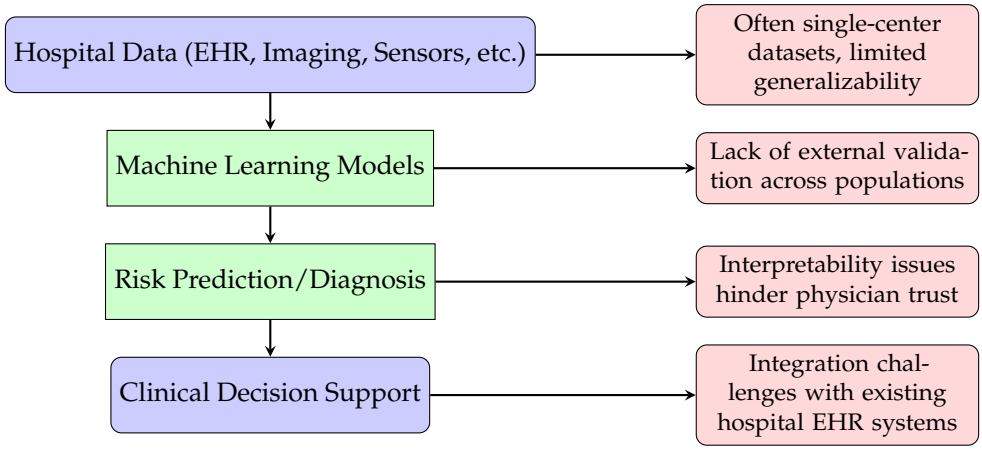

Figure 10. Workflow of ML-based DVT applications in clinical settings, along with common barriers.

7. Conclusions

The application of ML approaches for the early detection of DVT demonstrates substantial potential to enhance diagnostic precision, reducing both false negative and false positive rates compared to conventional approaches. These advances reflect a paradigm shift toward more accurate, efficient, and scalable diagnostic workflows.

Beyond early detection, ML models have proven effective in risk stratification and predictive analytics, identifying high-risk patients before the onset of clinical symptoms. By leveraging multimodal data—including imaging, laboratory findings, and electronic

health records—ML-driven systems can enable preventive patient-centered interventions that improve outcomes and support precision medicine initiatives.

Emerging applications, such as continuous monitoring through wearable sensors, further expand the roles of ML in clinical care by extending surveillance beyond hospital settings. These innovations hold promise for safer post-treatment follow-up, the early detection of complications, and the development of proactive care strategies.

Despite these advances, multiple barriers hinder the large-scale clinical adoption of ML. These include limited dataset availability, lack of standardized data collection and curation protocols, insufficient interpretability of complex models, and a scarcity of prospective multicenter clinical validation. Addressing these challenges will require interdisciplinary collaboration between data scientists, clinicians, and regulatory bodies to ensure the safe, transparent, and effective deployment of ML tools in healthcare systems.

In conclusion, ML represents a transformative approach to DVT diagnosis, risk assessment, and monitoring, offering the potential to optimize diagnostic workflows, improve patient care, and support the evolution of precision medicine. Future research should emphasize the development of interoperable datasets, robust external validation protocols, and interpretable ML frameworks to accelerate their clinical translation and maximize the impacts of these technologies in real-world healthcare environments.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/eng6090243/s1, Supplementary File S1 (SupplementaryMaterials.pdf) Figure S1. Distribution of the top 10 journals contributing to the initial article pool. Figure S2. Distribution of articles by author contributions. Figure S3. Distribution of articles by funding sources. Figure S4. Distribution of articles by type. Figure S5. VosViewer keyword co-occurrence map. Figure S6. Citation network among included articles. Supplementary File S2 (Risk_of_Bias.docx)—QUADAS Risk of Bias Assessment. A comprehensive evaluation of methodological quality across included studies conducted using the QUADAS checklist. Supplementary File S3 (PRIMA_Checklist.docx)—PRISMA Checklist. Completed PRISMA checklist providing a summary of adherence to systematic review reporting standards. Detailed information regarding journals, authors, and datasets included in this review is provided in the Appendices A–C (Tables A1–A3).

Author Contributions: Conceptualization, E.E.G.-G. and E.I.-G.; data curation, G.M.G.-A. and O.A.A.-C.; formal analysis, R.J.-R. and C.R.; funding acquisition, E.I.-G.; investigation, A.A.C.Z. and G.M.G.-A.; methodology, A.A.C.Z. and M.A.G.-G.; project administration, E.I.-G.; resources, E.E.G.-G.; software, A.A.C.Z. and O.A.A.-C.; supervision, E.E.G.-G. and E.I.-G.; validation, M.A.G.-G. and R.J.-R.; visualization, C.R. and O.A.A.-C.; writing—original draft, A.A.C.Z.; writing—review and editing, E.I.-G. and E.E.G.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Universidad Autónoma de Baja California (UABC) through the 25th internal call for research projects with grant number 402/6/C/53/25.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author(s).

Acknowledgments: The authors are thankful to the SECIHTI (Secretaría de Ciencia, Humanidades, Tecnología e Innovación) for the scholarship awarded to Andre Axel Cadena Zepeda and Marco Antonio Gómez-Guzmán.

Conflicts of Interest: The authors declare no conflicts of interest. The founding sponsors had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Table A1. Journals with one article.

Journal	Year
Annals of Vascular Surgery	2024
Clinical Nursing Research	2024
International Journal of Medical Informatics	2024
Journal of Thrombosis and Thrombolysis	2024
Journal of Orthopaedic Surgery	2024
Computers in Biology and Medicine	2024
Int. J. Gen. Med.	2024
Blood Adv.	2024
Thrombosis Research	2023
Clinical Spine Surgery	2023
Open Life Sciences	2023
Global Spine Journal	2023
Cancers	2023
J. Biomed. Inform.	2023
Journal of Obstetrics and Gynaecology	2023
Pulmonary Circulation	2022
Nature Medicine	2019
BMJ	2018
Cochrane Database Syst. Rev.	2018
ACM Computing Surveys	2018
arXiv preprint	2017
Aten. Primaria	2016
IEEE Transactions on Knowledge	2010
Circulation	2003

Appendix B

Table A2. Authors with one article.

Author	Year
Adler A	2024
Danilatou V	2024
Ma S	2024
Sheng Y	2024
Zhang L	2024
Zhou H	2024
Abraham J	2023
Chase H	2023
He W	2023
Karabacak M	2023
Li J	2023
Munoz A J	2023
Bao H	2021
Das R	2021
Ryan L	2021
Cohen AT	2020
Rezaee M	2020
Topol E J	2019
Sachdeva A	2018
Velickovic P	2018
Fuentes Camps E	2016
Gunderson C G	2014
Pan S J	2010
White R H	2003
Vilalta R	2002

Appendix C

Table A3. Multiple datasets in one article.

Country/Source	Year
China–U.S.	2024
Germany–U.K.–Greece	2024
China–Italy	2023
China–Synthetic Data	2020
Multi-Institutional (U.S.)	2020
Synthetic + Real (Mexico)	2020

References

1. Fuentes Camps, E.; Luis del Val García, J.; Bellmunt Montoya, S.; Hmimina Hmimina, S.; Gómez Jabalera, E.; Muñoz Pérez, M.Á. Estudio coste efectividad del proceso diagnóstico de la trombosis venosa profunda desde la atención primaria. *Aten. Primaria* **2016**, *48*, 251–257. [CrossRef]
2. Gunderson, C.G.; Chang, J.J. Overuse of compression ultrasound for patients with lower extremity cellulitis. *Thromb. Res.* **2014**, *134*, 846–850. [CrossRef]
3. Harder, E.M.; Desai, O.; Marshall, P.S. Clinical probability tools for deep venous thrombosis, pulmonary embolism, and bleeding. *Clin. Chest Med.* **2018**, *39*, 473–482. [CrossRef]
4. Sachdeva, A.; Dalton, M.; Lees, T. Graduated compression stockings for prevention of deep vein thrombosis. *Cochrane Database Syst. Rev.* **2018**, *11*, CD001484. [CrossRef]
5. Stubbs, M.J.; Mouyis, M.; Thomas, M. Deep vein thrombosis. *BMJ* **2018**, *360*, k351. [CrossRef]
6. Kraaijpoel, N.; Carrier, M.; Le Gal, G.; McInnes, M.D.F.; Salameh, J.P.; McGrath, T.A.; van Es, N.; Moher, D.; Büller, H.R.; Bossuyt, P.M.; et al. Diagnostic accuracy of three ultrasonography strategies for deep vein thrombosis of the lower extremity: A systematic review and meta-analysis. *PLoS ONE* **2020**, *15*, e0228788. [CrossRef] [PubMed]
7. White, R.H. The epidemiology of venous thromboembolism. *Circulation* **2003**, *107*, I4–I8. [CrossRef]
8. Lippi, G.; Mattiuzzi, C.; Franchini, M. Sleep apnea and venous thromboembolism: A systematic review. *Thromb. Haemost.* **2015**, *114*, 958–963. [CrossRef] [PubMed]
9. Fong-Mata, M.B.; García-Guerrero, E.E.; Mejía-Medina, D.A.; López-Bonilla, O.R.; Villarreal-Gómez, L.J.; Zamora-Arellano, F.; López-Mancilla, D.; Inzunza-González, E. An artificial neural network approach and a data augmentation algorithm to systematize the diagnosis of Deep-Vein Thrombosis by using Wells' criteria. *Electronics* **2020**, *9*, 1810. [CrossRef]
10. Contreras-Luján, E.E.; García-Guerrero, E.E.; López-Bonilla, O.R.; Tlelo-Cuautle, E.; López-Mancilla, D.; Inzunza-González, E. Evaluation of machine learning algorithms for early diagnosis of deep venous thrombosis. *Math. Comput. Appl.* **2022**, *27*, 24. [CrossRef]
11. Danilatou, V.; Dimopoulos, D.; Kostoulas, T.; Douketis, J. Machine learning-based predictive models for patients with venous thromboembolism: A systematic review. *Thromb. Haemost.* **2024**, *114* , 958–963. [CrossRef] [PubMed]
12. Oppenheimer, J.; Mandegaran, R.; Staabs, F.; Adler, A.; Singöhl, S.; Kainz, B.; Heinrich, M.; Geroulakos, G.; Spiliopoulos, S.; Avgerinos, E. Remote Expert DVT Triaging of Novice-User Compression Sonography with AI-Guidance. *Ann. Vasc. Surg.* **2024**, *99*, 272–279. [CrossRef]
13. Wang, K.Y.; Ikwuezunma, I.; Puvanesarajah, V.; Babu, J.; Margalit, A.; Raad, M.; Jain, A. Using Predictive Modeling and Supervised Machine Learning to Identify Patients at Risk for Venous Thromboembolism Following Posterior Lumbar Fusion. *Glob. Spine J.* **2023**, *13*, 1097–1103. [CrossRef]
14. Nakayama, Y.; Sato, M.; Okamoto, M.; Kondo, Y.; Tamura, M.; Minagawa, Y.; Uchiyama, M.; Horii, Y. Deep learning-based classification of adequate sonographic images for self-diagnosing deep vein thrombosis. *PLoS ONE* **2023**, *18*, e0282747. [CrossRef]
15. Chen, X.; Hou, M.; Wang, D. Machine learning-based model for prediction of deep vein thrombosis after gynecological laparoscopy: A retrospective cohort study. *Medicine* **2024**, *103*, e36717. [CrossRef] [PubMed]
16. Wei, C.; Wang, J.; Yu, P.; Li, A.; Xiong, Z.; Yuan, Z.; Yu, L.; Luo, J. Comparison of different machine learning classification models for predicting deep vein thrombosis in lower extremity fractures. *Sci. Rep.* **2024**, *14*, 6901 . [CrossRef]
17. Nafee, T.; Gibson, C.M.; Travis, R.; Yee, M.K.; Kerneis, M.; Chi, G.; AlKhalfan, F.; Hernandez, A.F.; Hull, R.D.; Cohen, A.T.; et al. Machine learning to predict venous thrombosis in acutely ill medical patients. *Res. Pract. Thromb. Haemost.* **2020**, *4*, 230–237. [CrossRef]
18. Nothnagel, K.; Aslam, M.F. Evaluating the benefits of machine learning for diagnosing deep vein thrombosis compared to gold standard ultrasound- a feasibility study. *BJGP Open* **2024**, *8* , BJGPO.2024.0057. [CrossRef]

19. Arun, R.; Joseph, B.R.C.; Muthukumar, B.; Ahilan, A. Deep vein thrombosis detection via combination of neural networks. *Biomed. Signal Process. Control* **2025**, *100*, 106972. [CrossRef]
20. Hwang, J.H.; Seo, J.W.; Kim, J.H.; Park, S.; Kim, Y.J.; Kim, K.G. Comparison between deep learning and conventional machine learning in classifying iliofemoral deep venous thrombosis upon CT venography. *Diagnostics* **2022**, *12*, 274. [CrossRef] [PubMed]
21. Zhou, H.; Jin, Y.; Chen, G.; Jin, X.; Chen, J.; Wang, J. Predictive modeling of lower extreme deep vein thrombosis following radical gastrectomy for gastric cancer: Based on multiple machine learning methods. *Sci. Rep.* **2024**, *14*, 15711 . [CrossRef]
22. Shohat, N.; Ludwick, L.; Sherman, M.B.; Fillingham, Y.; Parvizi, J. Using machine learning to predict venous thromboembolism and major bleeding events following total joint arthroplasty. *Sci. Rep.* **2023**, *13*, 2197 . [CrossRef]
23. Abraham, J.; Bartek, B.; Meng, A.; Ryan King, C.; Xue, B.; Lu, C.; Avidan, M.S. Integrating machine learning predictions for perioperative risk management: Towards an empirical design of a flexible-standardized risk assessment tool. *J. Biomed. Inform.* **2023**, *137*, 104270. [CrossRef] [PubMed]
24. Wells, P. Predictive analytics by deep machine learning: A call for next-gen tools to improve health care. *Res. Pract. Thromb. Haemost.* **2020**, *4*, 181–182. [CrossRef]
25. Jin, S.; Qin, D.; Liang, B.S.; Zhang, L.C.; Wei, X.X.; Wang, Y.J.; Zhuang, B.; Zhang, T.; Yang, Z.P.; Cao, Y.W.; et al. Machine learning predicts cancer-associated deep vein thrombosis using clinically available variables. *Int. J. Med. Inform.* **2022**, *161*, 104733 . [CrossRef]
26. Liu, S.; Zhang, F.; Xie, L.; Wang, Y.; Xiang, Q.; Yue, Z.; Feng, Y.; Yang, Y.; Li, J.; Luo, L.; et al. Machine learning approaches for risk assessment of peripherally inserted Central catheter-related vein thrombosis in hospitalized patients with cancer. *Int. J. Med. Inform.* **2019**, *129*, 175–183. [CrossRef]
27. Sabra, S.; Mahmood Malik, K.; Alobaidi, M. Prediction of venous thromboembolism using semantic and sentiment analyses of clinical narratives. *Comput. Biol. Med.* **2018**, *94*, 1–10. [CrossRef]
28. Guo, X.; Xu, H.; Zhang, J.; Hao, B.; Yang, T. A systematic review and meta-analysis of risk prediction models for post-thrombotic syndrome in patients with deep vein thrombosis. *Heliyon* **2023**, *9*, e22226. [CrossRef]
29. Kang, J.W.; Kim, K.T.; Park, J.W.; Lee, S.J. Classification of deep vein thrombosis stages using convolutional neural network of electromyogram with vibrotactile stimulation toward developing an early diagnostic tool: A preliminary study on a pig model. *PLoS ONE* **2023**, *18*, e0281219. [CrossRef] [PubMed]
30. Rezaee, M.; Putrenko, I.; Takeh, A.; Ganna, A.; Ingelsson, E. Development and validation of risk prediction models for multiple cardiovascular diseases and Type 2 diabetes. *PLoS ONE* **2020**, *15*, e0235758. [CrossRef] [PubMed]
31. Silva, L.O.d.; Silva, M.C.B.d.; Ribeiro, G.A.S.; Camargo, T.F.O.d.; Santos, P.V.D.; Mendes, G.d.S.; Paiva, J.P.Q.d.; Soares, A.d.S.; Reis, M.R.d.C.; Loureiro, R.M.; et al. Artificial intelligence-based pulmonary embolism classification: Development and validation using real-world data. *PLoS ONE* **2024**, *19*, e0305839. [CrossRef] [PubMed]
32. Liu, S.H.; Wang, J.J.; Chen, W.; Pan, K.L.; Su, C.H. An examination system to detect deep vein thrombosis of a lower limb using light reflection rheography. *Sensors* **2021**, *21*, 2446. [CrossRef]
33. Hou, T.; Qiao, W.; Song, S.; Guan, Y.; Zhu, C.; Yang, Q.; Gu, Q.; Sun, L.; Liu, S. The use of machine learning techniques to predict deep vein thrombosis in rehabilitation inpatients. *Clin. Appl. Thromb. Hemost.* **2023**, *29*, 10760296231179438. [CrossRef] [PubMed]
34. Kainz, B.; Heinrich, M.P.; Makropoulos, A.; Oppenheimer, J.; Mandegaran, R.; Sankar, S.; Deane, C.; Mischkewitz, S.; Al-Noor, F.; Rawdin, A.C.; et al. Non-invasive diagnosis of deep vein thrombosis from ultrasound imaging with machine learning. *npj Digit. Med.* **2021**, *4*, 137. [CrossRef]
35. Wang, X.; Xi, H.; Geng, X.; Li, Y.; Zhao, M.; Li, F.; Li, Z.; Ji, H.; Tian, H. Artificial Intelligence-Based Prediction of Lower Extremity Deep Vein Thrombosis Risk After Knee/Hip Arthroplasty. *Clin. Appl. Thromb./Hemost.* **2023**, *29*, 1–9 . [CrossRef]
36. Zhang, J.; Chen, J.; Yang, X.; Han, J.; Chen, X.; Fan, Y.; Zheng, H. Novel risk prediction models, involving coagulation, thromboelastography, stress response, and immune function indicators, for deep vein thrombosis after radical resection of cervical cancer and ovarian cancer. *J. Obstet. Gynaecol.* **2023**, *43*, 2204162 . [CrossRef]
37. Yu, T.; Shen, R.; You, G.; Lv, L.; Kang, S.; Wang, X.; Xu, J.; Zhu, D.; Xia, Z.; Zheng, J.; et al. Machine learning-based prediction of the post-thrombotic syndrome: Model development and validation study. *Front. Cardiovasc. Med.* **2022**, *9*, 990788 . [CrossRef] [PubMed]
38. Xu, Q.; Lei, H.; Li, X.; Li, F.; Shi, H.; Wang, G.; Sun, A.; Wang, Y.; Peng, B. Machine learning predicts cancer-associated venous thromboembolism using clinically available variables in gastric cancer patients. *Heliyon* **2023**, *9*, e12681. [CrossRef]
39. Zhang, L.; Yu, R.; Chen, K.; Zhang, Y.; Li, Q.; Chen, Y. Enhancing deep vein thrombosis prediction in patients with coronavirus disease 2019 using improved machine learning model. *Comput. Biol. Med.* **2024**, *173*, 108294 . [CrossRef]
40. Jin, J.; Lu, J.; Su, X.; Xiong, Y.; Ma, S.; Kong, Y.; Xu, H. Development and validation of an ICU-venous thromboembolism prediction model using machine learning approaches: A multicenter study. *Int. J. Gen. Med.* **2024**, *17*, 3279–3292. [CrossRef]

41. Qiao, N.; Zhang, Q.; Chen, L.; He, W.; Ma, Z.; Ye, Z.; He, M.; Zhang, Z.; Zhou, X.; Shen, M.; et al. Machine learning prediction of venous thromboembolism after surgeries of major sellar region tumors. *Thromb. Res.* **2023**, *226*, 1–8. [CrossRef]
42. Wu, X.; Wang, Z.; Zheng, L.; Yang, Y.; Shi, W.; Wang, J.; Liu, D.; Zhang, Y. Construction and verification of a machine learning-based prediction model of deep vein thrombosis formation after spinal surgery. *Int. J. Med. Inform.* **2024**, *192*, 105609 . [CrossRef] [PubMed]
43. Wei, Y.; Lin, Q.; Hu, J. Study and Validation Protocol of Risk Prediction Model for Deep Venous Thrombosis After Severe Traumatic Brain Injury Based on Machine Learning Algorithms. In *Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2024; Volume 14710, pp. 328–342. [CrossRef]
44. Zhang, J.; Shao, Y.; Zhou, H.; Li, R.; Xu, J.; Xiao, Z.; Lu, L.; Cai, L. Prediction model of deep vein thrombosis risk after lower extremity orthopedic surgery. *Heliyon* **2024**, *10*, e29517. [CrossRef] [PubMed]
45. Lam, B.D.; Chrysafi, P.; Chiasakul, T.; Khosla, H.; Karagkouni, D.; McNichol, M.; Adamski, A.; Reyes, N.; Abe, K.; Mantha, S.; et al. Machine learning natural language processing for identifying venous thromboembolism: Systematic review and meta-analysis. *Blood Adv.* **2024**, *8*, 2991–3000. [CrossRef] [PubMed]
46. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph Attention Networks. *arXiv* **2018**, arXiv:1710.10903. [CrossRef]
47. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]
48. Vilalta, R.; Drissi, Y. A Perspective View and Survey of Meta-Learning. *Artif. Intell. Rev.* **2002**, *18*, 77–95. [CrossRef]
49. Sheller, M.J.; Edwards, B.; Reina, G.A.; Martin, J.; Pati, S.; Kotrotsou, A.; Milchenko, M.; Xu, W.; Marcus, D.; Colen, R.R.; et al. Federated Learning in Medicine: Facilitating Multi-Institutional Collaborations without Sharing Patient Data. *Sci. Rep.* **2020**, *10*, 12598. [CrossRef]
50. Tjoa, E.; Guan, C. Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 4793–4813. [CrossRef]
51. Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A Survey of Methods for Explaining Black Box Models. *ACM Comput. Surv.* **2018**, *51*, 1–42. [CrossRef]
52. Esteban, C.; Hyland, S.L.; Rätsch, G. Real-valued (Medical) Time Series Generation with Recurrent Conditional GANs. *arXiv* **2017**, arXiv:1706.02633. [CrossRef]
53. Sun, Y.; Bao, H.; Sun, W.; Liu, H. The Role of Digital Twins in Personalized Healthcare: Promises and Challenges. *J. Pers. Med.* **2021**, *11*, 745. [CrossRef]
54. Topol, E.J. High-performance medicine: The convergence of human and artificial intelligence. *Nat. Med.* **2019**, *25*, 44–56. [CrossRef]
55. Zeng, Y.; Chen, Y.; Zhu, D.; Xu, J.; Zhang, X.; Ying, H.; Song, X.; Zhou, R.; Wang, Y.; Yu, F. Machine learning assisted radiomics in predicting postoperative occurrence of deep venous thrombosis in patients with gastric cancer. *BMC Cancer* **2025**, *25*, 220. [CrossRef] [PubMed]
56. Ye, M.; Liu, C.; Yang, D.; Gao, H. Development and validation of a risk prediction model for acute kidney injury in coronary artery disease. *BMC Cardiovasc. Disord.* **2025**, *25*, 12. [CrossRef] [PubMed]
57. Iding, A.F.J.; Ten Cate, V.; Ten Cate, H.; Wild, P.S.; Ten Cate-Hoek, A.J. Untangling profiles of postthrombotic syndrome using unsupervised machine learning. *Blood Adv.* **2025**, *9*, 3631–3641. [CrossRef]
58. Dalil, D.; Esmaeili, S.; Safaee, E.; Asgari, S.; Kejani, N. The prediction of venous thromboembolism using artificial intelligence and machine learning in lower extremity arthroplasty: A Systematic Review. *Arthroplast. Today* **2025**, *33*, 101672. [CrossRef] [PubMed]
59. Yalzadeh, D.; Cho, N.Y.; Tabibian, D.; Song, J.; Cherif, A.; Badiee, B.; Chaturvedi, A.; Singer, G.; Benharash, P. Comparison of frailty measures in predicting outcomes after emergency general surgery. *Surgery* **2025**, *182*, 109317. [CrossRef]
60. Muñoz Martín, A.J.; Lecumberri, R.; Souto, J.C.; Obispo, B.; Sanchez, A.; Aparicio, J.; Aguayo, C.; Gutierrez, D.; García Palomo, A.; Benavent, D.; et al. Prediction model for major bleeding in anticoagulated patients with cancer-associated venous thromboembolism using machine learning and natural language processing. *Clin. Transl. Oncol.* **2025**, *27*, 1816–1825. [CrossRef]
61. Shen, J.; Xue, B.; Kannampallil, T.; Lu, C.; Abraham, J. A novel generative multi-task representation learning approach for predicting postoperative complications in cardiac surgery patients. *J. Am. Med. Inform. Assoc.* **2025**, *32*, 459–469. [CrossRef]
62. Fu, M.; Li, X.; Wang, Z.; Yang, Q.; Yu, G. Development and validation of machine learning-based prediction model for central venous access device-related thrombosis in children. *Thromb. Res.* **2025**, *247*, 109276. [CrossRef] [PubMed]
63. Huang, Y.; Liang, H.; Huang, S.; Xie, X.; Deng, B.; Liang, W. Evaluation of risk factors for thromboembolic events in multiple myeloma patients using multiple machine learning models. *Medicine* **2025**, *104*, e41428. [CrossRef] [PubMed]
64. Ma, G.; Chen, S.; Peng, S.; Yao, N.; Hu, J.; Xu, L.; Chen, T.; Wang, J.; Huang, X.; Zhang, J. Construction and validation of a nomogram prediction model for the catheter-related thrombosis risk of central venous access devices in patients with cancer: A prospective machine learning study. *J. Thromb. Thrombolysis* **2025**, *58*, 220–231. [CrossRef]

65. Xu, L.; Da, M. Incidence and risk factors of lower limb deep vein thrombosis in psychiatric inpatients by applying machine learning to electronic health records: A retrospective cohort study. *Clin. Epidemiol.* **2025**, *17*, 197–209. [CrossRef]
66. Tian, Y.; Liu, J.; Wu, S.; Zheng, Y.; Han, R.; Bao, Q.; Li, L.; Yang, T. Development and validation of a deep learning-enhanced prediction model for the likelihood of pulmonary embolism. *Front. Med.* **2025**, *12*, 1506363. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.